

Extracting Sub-Exposure Images from a Single Capture Through Fourier-based Optical Modulation

Shah Rez Khan¹, Martin Feldman², and Bahadır K. Gunturk¹

¹Dept. of Electrical and Electronics Engineering, Istanbul Medipol University, Istanbul, Turkey

²Div. of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA

Abstract—Through pixel-wise optical coding of images during exposure time, it is possible to extract sub-exposure images from a single capture. Such a capability can be used for different purposes, including high-speed imaging, high-dynamic-range imaging and compressed sensing. In this paper, we demonstrate a sub-exposure image extraction method, where the exposure coding pattern is inspired from frequency division multiplexing idea of communication systems. The coding masks modulate sub-exposure images in such a way that they are placed in non-overlapping regions in Fourier domain. The sub-exposure image extraction process involves digital filtering of the captured signal with proper band-pass filters. The prototype imaging system incorporates a Liquid Crystal over Silicon (LCoS) based spatial light modulator synchronized with a camera for pixel-wise exposure coding.

I. INTRODUCTION

Coded aperture and coded exposure photography methods, which involve control of aperture shape and exposure pattern during exposure period, present new capabilities and advantages over traditional photography. In coded aperture photography, the aperture shape is designed to achieve certain goals. For example, the aperture shape can be designed to improve depth estimation accuracy as a part of depth-from-defocus technique [1], or to improve deblurring performance through adjusting the zero crossings of point spread function [2]. Coded aperture photography may involve capture of multiple images, where each image is captured with a different aperture shape, for instance, to acquire light field [3], or to improve depth estimation and deblurring performance [4]. Using coded aperture, it is possible to do lensless imaging as well [5], [6].

In coded exposure photography, the exposure pattern is controlled during exposure period. The coding can be global, where all pixels are exposed together with a temporal pattern, or pixel-wise, where each pixel has its own exposure pattern. An example of global exposure coding is the flutter shutter technique [7], where the shutter is opened and closed according to a specific pattern during exposure period to enable better recovery from motion blur. The flutter shutter idea can also be used for high-speed imaging [8]. Pixel-wise exposure control presents more flexibility and wider range of applications compared to global exposure coding. An example of pixel-wise exposure control is presented in [9], where the

goal is to spatially adapt the dynamic range of the captured image. Pixel-wise coded exposure imaging can also be used for focal stacking through moving the lens during the exposure period [10], and for high-dynamic-range video through per-pixel exposure offsets [11].

Pixel-wise exposure control is also used for high-speed imaging by extracting sub-exposure images from a single capture. In [12], pixels are exposed according to a regular non-overlapping pattern on the space-time exposure grid. Some spatial samples are skipped in one time period to take samples in another time period to improve temporal resolution. In other words, spatial resolution is traded off for temporal resolution. In [13], there is also a non-overlapping space-time exposure sampling pattern; however, unlike the global spatial-temporal resolution trade-off approach of [12], the samples are integrated in various ways to spatially adapt spatial and temporal resolutions according to the local motion. For fast moving regions, fine temporal sampling is preferred; for slow moving regions, fine spatial sampling is preferred. Instead of a regular exposure pattern, random patterns can also be used [14], [15], [16]. In [14], pixel-wise random sub-exposure masks are used during exposure period. The reconstruction algorithm utilizes the spatial correlation of natural images and the brightness constancy assumption in temporal domain to achieve high-speed imaging. In [15], [16], the reconstruction algorithm is based on sparse dictionary learning. While learning-based approaches may yield outstanding performance, one drawback is that the dictionaries need to be re-trained each time a related parameter, such as target frame rate, is changed.

Alternative to arbitrary pixel-wise exposure patterns, it is also proposed to have row-wise control [17] and translated exposure mask [18], [19]. In [17], instead of traditional rolling shutter, row-wise coded exposure pattern is proposed. The row-wise exposure pattern can be designed to achieve high-speed imaging, high-dynamic-range imaging, and adaptive auto exposure. In [18], binary transmission masks are translated during exposure period for exposure coding; sub-exposure images are then reconstructed using an alternating projections algorithm. The same coding scheme is also used in [19], but a different image reconstruction approach is taken.

Pixel-wise exposure control can be implemented using regular image sensors with the help of additional optical elements. In [9], an LCD panel is placed in front of a camera to spatially control light attenuation. With such a system, pixel-wise exposure control is difficult since the LCD attenuator

is optically defocused. In [20], an LCD panel is placed on the intermediate image plane, which allows better pixel-by-pixel exposure control. One disadvantage of using transmissive LCD panels is the low fill factor due to drive circuit elements between the liquid crystal elements. In [21], a DMD reflector is placed on the intermediate image plane. DMD reflectors have high fill factor and high contrast ratio, thus they can produce sharper and higher dynamic range images compared to LCD panels. One drawback of the DMD based design is that the micromirrors on a DMD device reflect light at two small angles, thus the DMD plane and the sensor plane must be properly inclined, resulting in “keystone” perspective distortion. That is, a square DMD pixel is imaged as a trapezoid shape on the sensor plane. As a result, pixel-to-pixel mapping between the DMD and the sensor is difficult. In [22], a reflective LCoS spatial light modulator (SLM) is used on the intermediate image plane. Because the drive circuits on an LCoS device are on the back, high fill factor is possible as opposed to the transmissive LCD devices. Compared to a DMD, one-to-one pixel correspondence is easier with an LCoS SLM; however, the light efficiency is not as good as the DMD approach. In [18], a lithographically patterned chrome-on-quartz binary transmission mask is placed on the intermediate image plane, and moved during exposure period with a piezoelectric stage for optical coding. This approach is limited in terms of the exposure pattern that can be applied.

In this paper, we demonstrate a sub-exposure image extraction idea. The idea, which is called frequency division multiplexed imaging (FDMI), was presented by Gunturk and Feldman as a conference paper [23]. While the FDMI idea was demonstrated by merging two separate images with a patterned glass based and an LCD panel based modulation in [23], it was not demonstrated for sub-exposure image extraction. Here, we apply the FDMI idea to extract sub-exposure images using an optical setup incorporating an LCoS SLM synchronized with a camera for exposure coding.

In Section 2, we present the problem of extracting sub-exposure images through space-time exposure coding, and review the FDMI approach. In Section 3, we present the optical setup used in the experiments. In Section 4, we provide experimental results with several coded image captures. In Section 5, we conclude the paper with some future research directions.

II. EXTRACTING SUB-EXPOSURE IMAGES FROM A SINGLE CAPTURE

There are various exposure coding schemes designed for extracting sub-exposure images from a single capture. First, we would like to present a formulation of the coding process, and then review the FDMI idea.

A. Space-Time Exposure Coding

Space-time exposure coding of an image can be formulated using a spatio-temporal video signal $I(x, y, t)$, where (x, y) are the spatial coordinates and t is the time coordinate. This

signal is modulated during an exposure period T with a mask $m(x, y, t)$ to generate an image:

$$I(x, y) = \int_0^T m(x, y, t) I(x, y, t) dt. \quad (1)$$

The mask $m(x, y, t)$ can be divided in time into a set of constant sub-exposure masks: $m_1(x, y)$ for $t \in (t_0, t_1)$, $m_2(x, y)$ for $t \in (t_1, t_2)$, ..., $m_N(x, y)$ for $t \in (t_{N-1}, t_N)$, where N is the number of masks, and t_0 and t_N are the start and end times of the exposure period. Incorporating the sub-exposure masks $m_i(x, y)$ into equation (1), the captured image becomes

$$I(x, y) = \sum_{i=1}^N m_i(x, y) \int_{t_{i-1}}^{t_i} I(x, y, t) dt = \sum_{i=1}^N m_i(x, y) I_i(x, y), \quad (2)$$

where we define the sub-exposure image $I_i(x, y) = \int_{t_{i-1}}^{t_i} I(x, y, t) dt$. The above equation states that the sub-exposure images $I_i(x, y)$ are modulated with the masks $m_i(x, y)$ and added up to form the recorded image $I(x, y)$. The goal of sub-exposure image extraction is to estimate the images $I_i(x, y)$ given the masks and the recorded image.

As we have already mentioned in Section 1, the reconstruction process might be based on different techniques, varying from simple interpolation [12] to dictionary learning [16]. The masks $m_i(x, y)$ can be chosen in different ways as well. Some of the masks, including non-overlapping uniform grid [12], coded rolling shutter [17], pixel-wise random exposure [15], and frequency division multiplexed imaging exposure [23], are illustrated in Figure 1.

B. Frequency Division Multiplexed Imaging

The frequency division multiplexed imaging (FDMI) idea [23] is inspired from frequency division multiplexing method in communication systems, where the communication channel is divided into non-overlapping subbands, each of which carry independent signals that are properly modulated. In case of FDMI, sub-exposure images are modulated such that they are placed in different regions in Fourier domain. By ensuring that the Fourier components of different sub-exposure images do not overlap, each sub-exposure image can be extracted from the captured signal through band-pass filtering. The FDMI idea is illustrated for two images in Figure 2. Two band-limited sub-exposure images $I_1(x, y)$ and $I_2(x, y)$ are modulated with horizontal and vertical sinusoidal masks $m_1(x, y)$ and $m_2(x, y)$ during exposure period. The masks can be chosen as raised cosines: $m_1(x, y) = a + b \cos(2\pi u_0 x)$ and $m_2(x, y) = a + b \cos(2\pi v_0 y)$, where a and b are positive constants with the condition $a \geq b$ so that the masks are non-negative, that is, optically realizable, and u_0 and v_0 are the spatial frequencies of the masks. The imaging system captures sum of the modulated images: $I(x, y) = m_1(x, y) I_1(x, y) + m_2(x, y) I_2(x, y)$.

The imaging process from the Fourier domain perspective is also illustrated in Figure 2. $\hat{I}_1(u, v)$ and $\hat{I}_2(u, v)$ are the Fourier transforms of the sub-exposure images $I_1(x, y)$ and $I_2(x, y)$, which are assumed to be band-limited. In Figure 2, the Fourier transforms $\hat{I}_1(u, v)$ and $\hat{I}_2(u, v)$ are shown

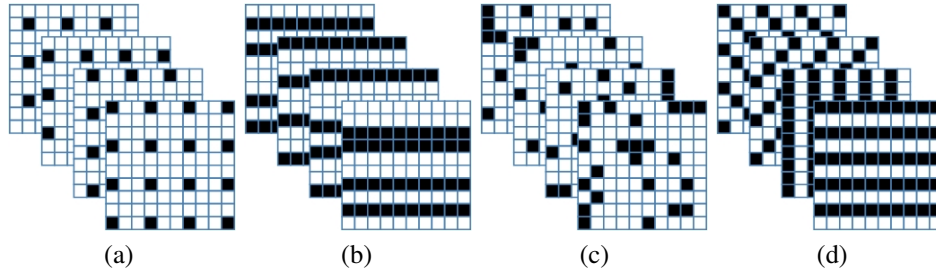


Fig. 1: Illustration of different exposure masks. (a) Non-overlapping uniform grid exposure [13], (b) Coded rolling shutter [17], (c) Pixel-wise random exposure [15], (d) Frequency division multiplexed imaging exposure [23].

as circular regions. The Fourier transforms of the masks are impulses: $\hat{M}_1(u, v) = a\delta(u, v) + (b/2)(\delta(u - u_0, v) + \delta(u + u_0, v))$ and $\hat{M}_2(u, v) = a\delta(u, v) + (b/2)(\delta(u, v - v_0) + \delta(u, v + v_0))$. As a result, the Fourier transform $\hat{I}(u, v)$ of the recorded image $I(x, y)$ includes $\hat{I}_1(u, v)$ and $\hat{I}_2(u, v)$ in its sidebands, and $\hat{I}_1(u, v) + \hat{I}_2(u, v)$ in the baseband. From the sidebands, the individual sub-exposure images can be recovered with proper band-pass filtering. From the baseband, the full-exposure image $I_1(x, y) + I_2(x, y)$ can be recovered with a low-pass filter.

It is possible to use other periodic signals instead of a cosine wave. For example, when a square wave is used, the baseband is modulated to all harmonics of the main frequency, where the weights of the harmonics decrease with a sinc function. Again, by applying band-pass filters on the first harmonics, sub-exposure images can be recovered.

III. PROTOTYPE DESIGN

The prototype system is based on the pixel-wise exposure control design that is adopted in several papers [22], [15], [16]. As shown in Figure 3, the system consists of an objective lens, three relay lenses, one polarizing beam splitter, an LCoS SLM, and a camera. Relay lenses are 100mm aspherical doublets; the objective lens is a 75mm aspherical doublet; the SLM is a Forth Dimension Display SXGA-3DM with 1280x1024 pixel resolution; and the camera is a Thorlabs DCU-224M monochromatic camera with 1280x1024 pixel resolution. The objective lens forms an image on an intermediate image plane, which is relayed onto the SLM. The SLM controls the exposure pattern of each pixel through changing the polarization states. The image reflected from the SLM is recorded by the camera. The camera and the SLM are synchronized using the trigger signal from the SLM.

IV. EXPERIMENTAL RESULTS

We conducted several experiments to demonstrate sub-exposure image extraction with FDMI. In each case, the exposure time of the camera is set to 45 milliseconds. In the first 30 millisecond period, the SLM pattern is a square wave in horizontal direction; in the second 15 millisecond period, the SLM pattern is a square wave in vertical direction. The square wave on the SLM is set to its highest possible spatial frequency; that is, one pixel black (no reflection) and pixel white (full reflection). Because of the pixel size difference,

one SLM pixel corresponds to an about 3×3 camera sensor pixel. It is, however, not necessary to detect/extract SLM pixels from the recorded image since the FDMI method only requires applying a band-pass filter to the recorded image.

In Figure 4, the target object is a printout, which is rotated as the camera captures an image. The image is exposure-coded as described above. The captured image is shown in Figure 4(a). A zoomed-in region is shown in Figure 4(b), where horizontal and vertical lines can be clearly seen. The Fourier transform of the image is shown in Figure 4(c). The band-pass filters to be applied on the sidebands and the low-pass filter to be applied on the baseband are also marked on the Fourier transform. By applying the band-pass filter to the horizontal sideband, the 30 millisecond sub-exposure image is extracted. By applying the band-pass filter to the vertical sideband, the 15 millisecond sub-exposure image is extracted. The baseband gives the sum of the sub-exposure images, as if there were no optical modulation.

In Figure 4(d), the baseband image is shown. The motion blur prevents seeing the details in the image, which has a full-exposure period of 45 milliseconds. The 30 millisecond sub-exposure image is shown in Figure 4(e); and the 15 millisecond sub-exposure image is shown in Figure 4(f). The readability of the text and the visibility of details are much improved in the 15 millisecond sub-exposure image.

In the second experiment, a ball is thrown in front of the camera. The image is recorded with the same exposure pattern. The recorded image, baseband image, and sideband sub-exposure images are shown in Figure 5. The 15 millisecond sub-exposure image is able to capture the ball shape with much less motion blur.

In the third experiment, shown in Figure 6, a planar surface with a text printed on is moved in front of the camera. The baseband is degraded with space varying motion blur. In the 15 millisecond sub-exposure image, the text becomes readable.

V. CONCLUSIONS

In this paper, we demonstrate a sub-exposure image extraction method, which is based on allocating different regions in Fourier domain to different sub-exposure images. We demonstrate the method for extracting two sub-exposure images. It is theoretically possible to extract more sub-exposure images by band-limiting the images and placing them compactly in

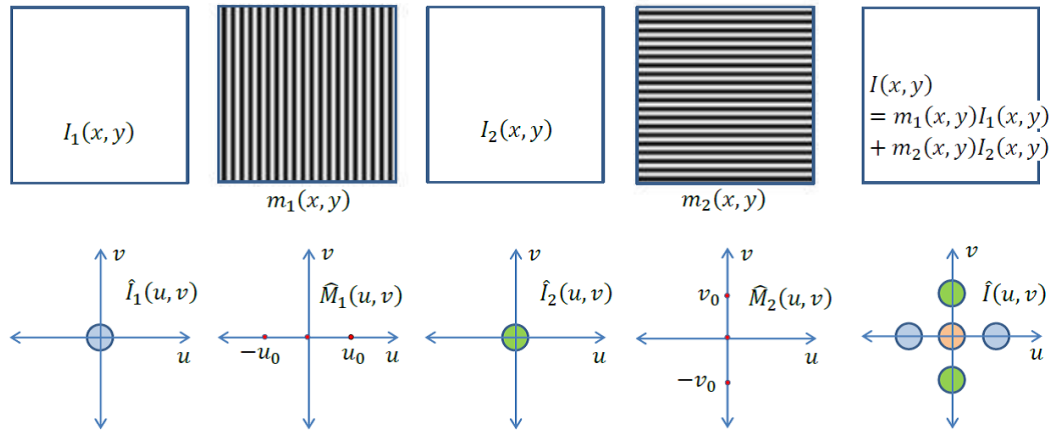


Fig. 2: Illustration of the FDMI idea with two images [23].

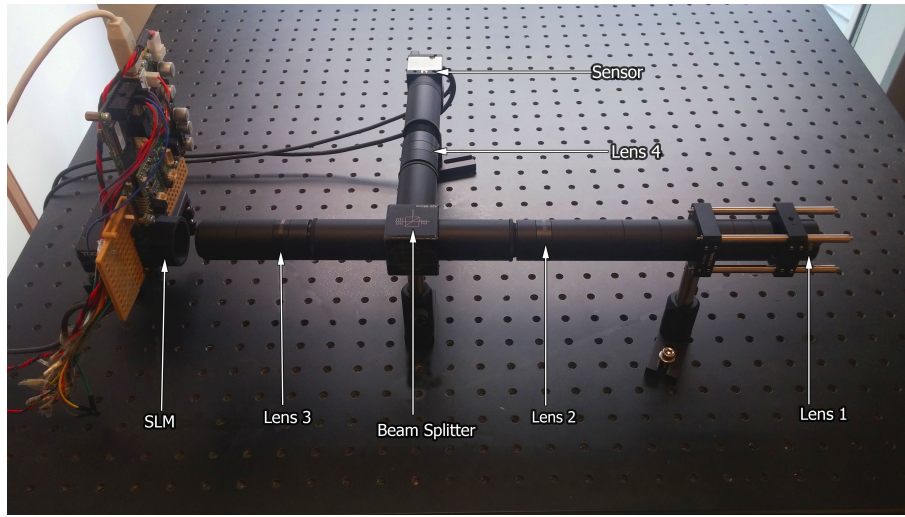
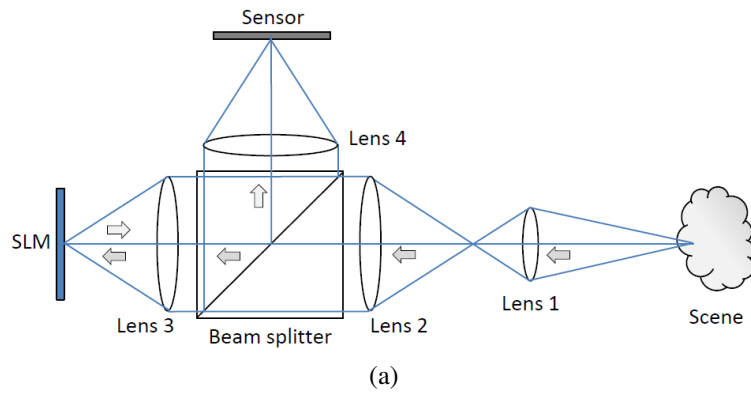


Fig. 3: Pixel-wise exposure control camera prototype. (a) Graphical illustration, (b) Actual prototype.

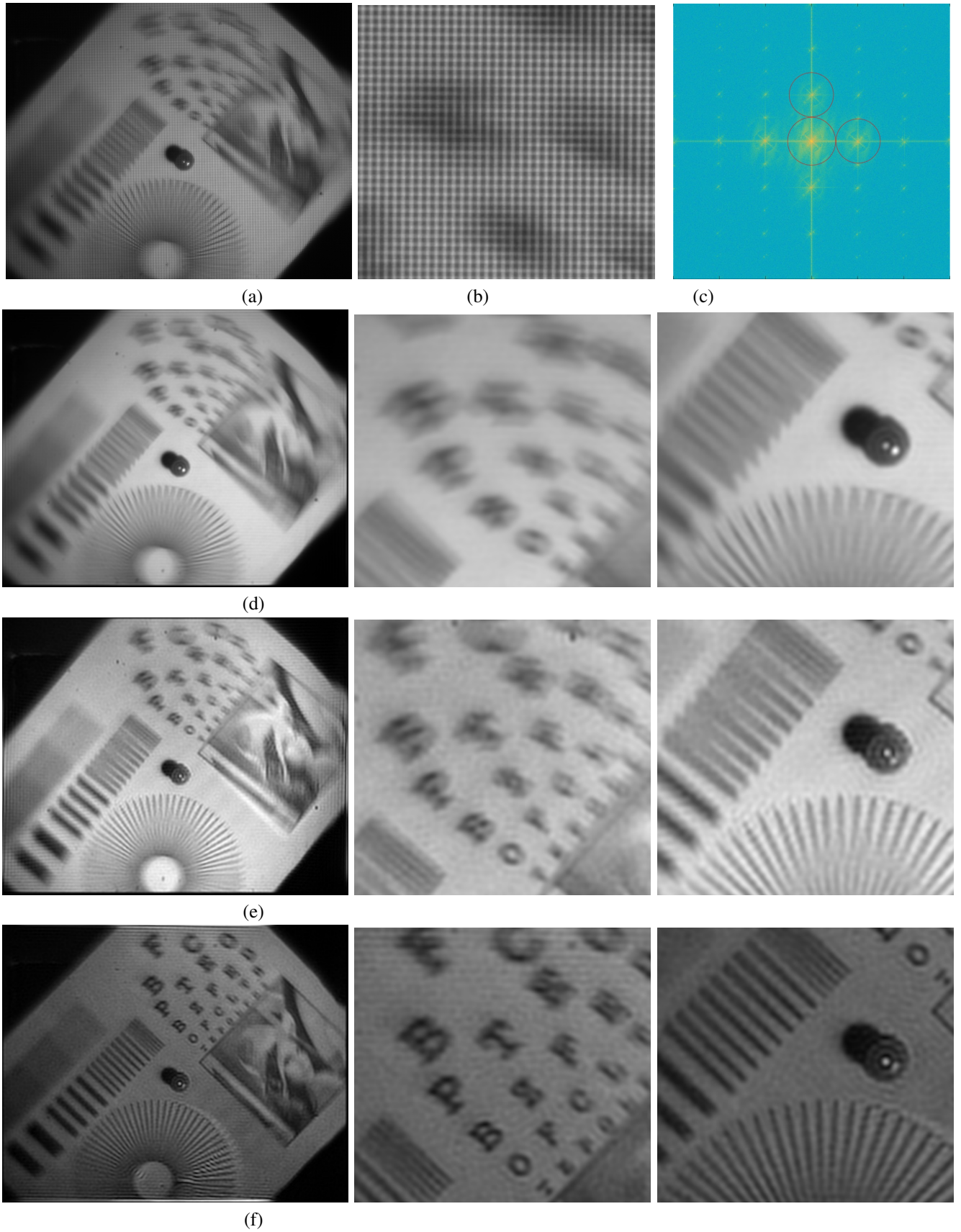


Fig. 4: Sub-exposure image extraction with FDMI. (a) Recorded exposure-coded image, (b) A zoomed-in region from (a), (c) Fourier transform (magnitude) of the image in (a), (d) Image extracted from the baseband, (e) Image extracted from the horizontal sideband, corresponding to 30 millisecond exposure period, (f) Image extracted from the vertical sideband, corresponding to 15 millisecond exposure period. Zoomed-in regions for (d), (e), and (f) are also shown.

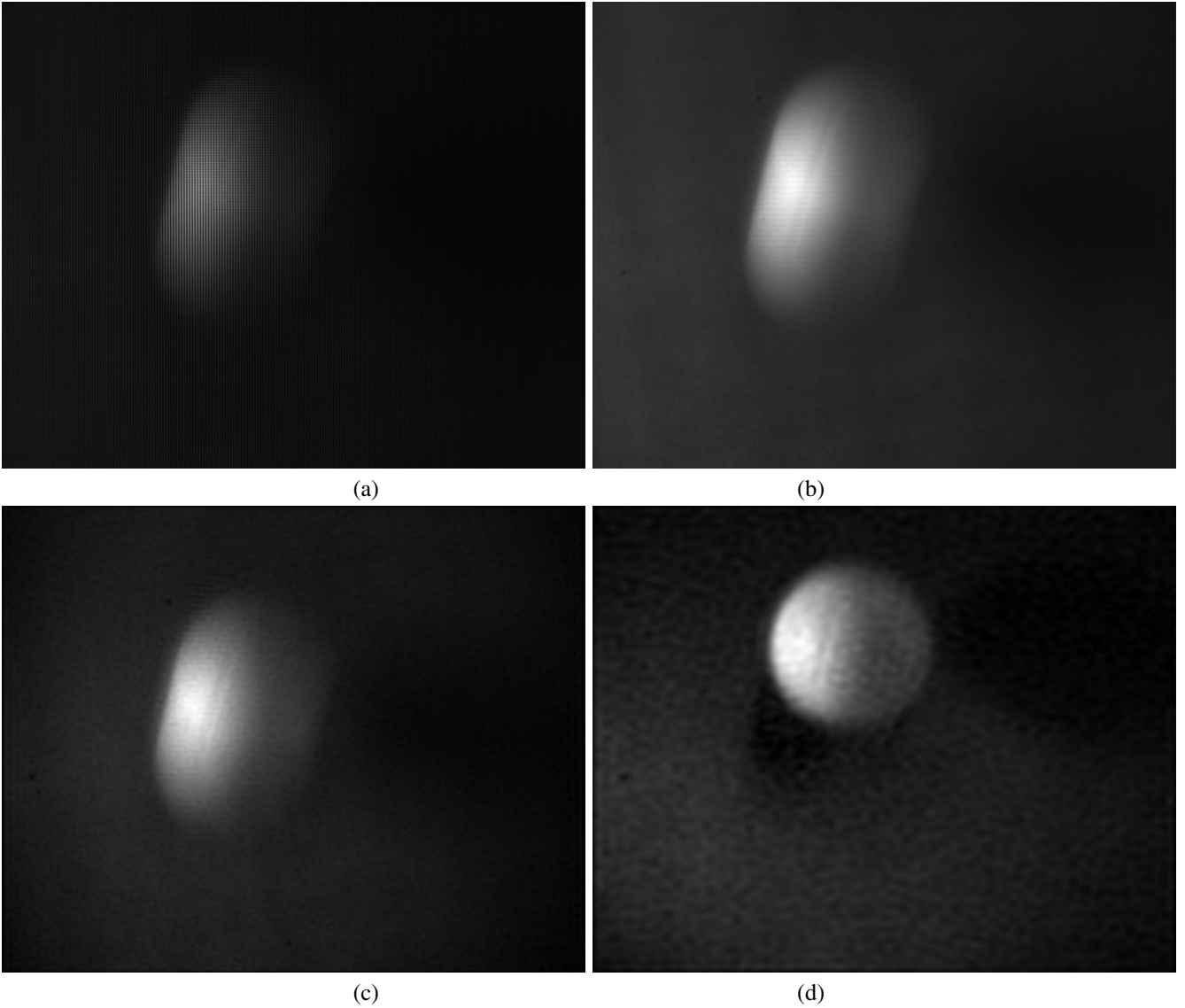


Fig. 5: Sub-exposure image extraction with FDMI. (a) Recorded exposure-coded image, (b) Baseband image, (c) 30 millisecond sub-exposure image, (d) 15 millisecond sub-exposure image.

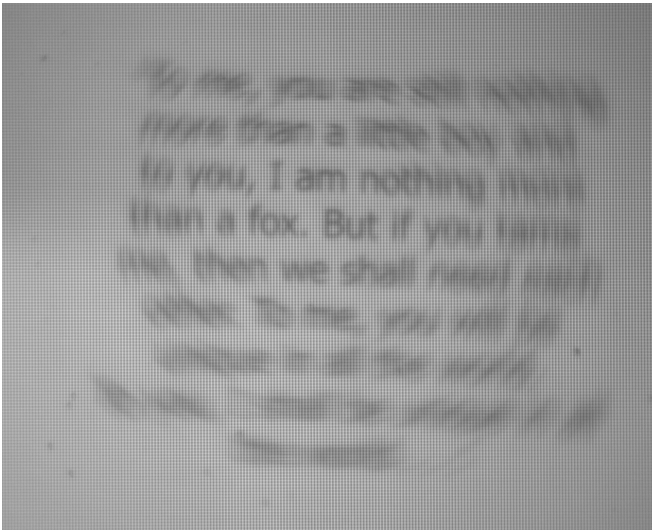
Fourier domain. This would require very accurate pixel-wise optical modulation, which may be possible by applying the exposure pattern directly on the sensor instead of using an SLM.

An advantage of the FDMI approach is the low computational complexity, which involves Fourier transform and band-pass filtering. The sub-exposure images from the sidebands and the full exposed image from the baseband are easily extracted by taking Fourier transform, band-pass filtering, and inverse Fourier transform.

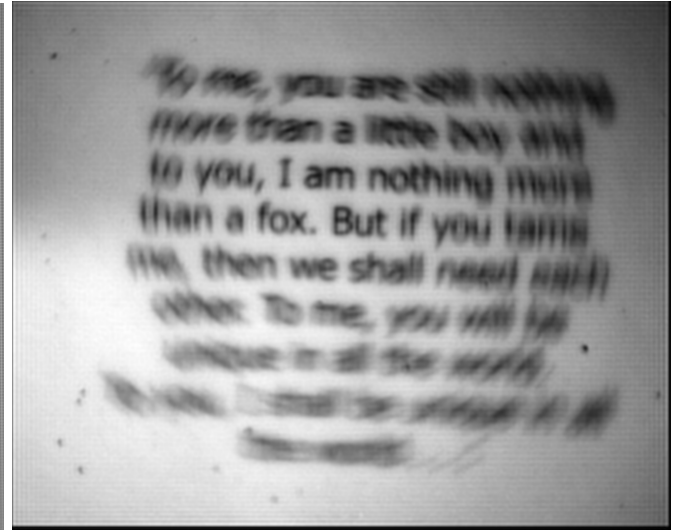
Extraction of sub-exposure images can be used for different purposes. In addition to high-speed imaging, one may try to estimate space-varying motion blur, which could be used in scene understanding, segmentation, and space-varying deblurring applications.

REFERENCES

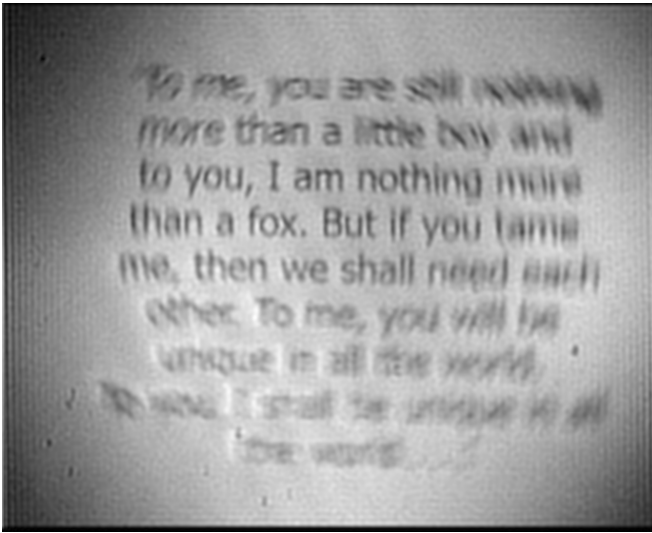
- [1] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Trans. on Graphics* **26**, Article No. 70 (2007).
- [2] C. Zhou and S. Nayar, "What are good apertures for defocus deblurring?" in "IEEE Int. Conf. on Computational Photography," (2009), pp. 1–8.
- [3] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. H. Chen, "Programmable aperture photography: multiplexed light field acquisition," in "ACM Trans. on Graphics," , vol. 27 (2008), vol. 27, p. Article No. 55.
- [4] C. Zhou, S. Lin, and S. K. Nayar, "Coded aperture pairs for depth from defocus and defocus deblurring," *Int. Journal of Computer Vision* **93**, 53–72 (2011).
- [5] Y. Sekikawa, S.-W. Leigh, and K. Suzuki, "Coded lens: Coded aperture for low-cost and versatile imaging," in "ACM SIGGRAPH," (2014).
- [6] M. S. Asif, A. Ayremlou, A. Veeraraghavan, and R. Baraniuk, "Flatcam: Replacing lenses with masks and computation," in "IEEE Int. Conf. on Computer Vision," (2015), pp. 12–15.



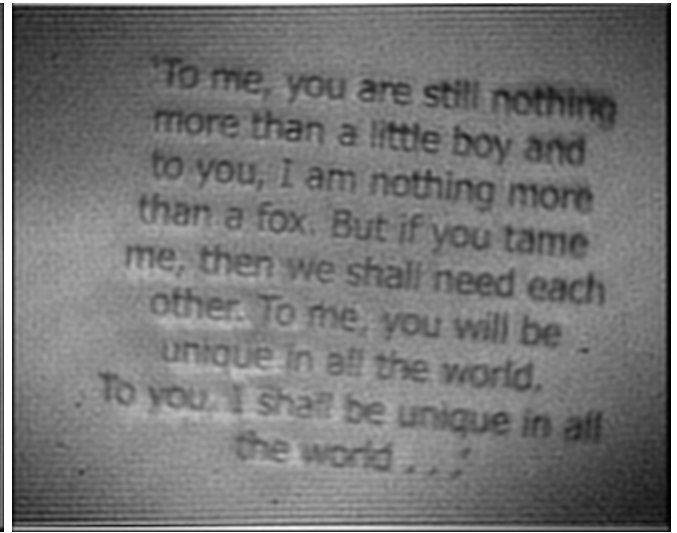
(a)



(b)



(c)



(d)

Fig. 6: Sub-exposure image extraction with FDMI. (a) Recorded exposure-coded image, (b) Baseband image, (c) 30 millisecond sub-exposure image, (d) 15 millisecond sub-exposure image.

- [7] R. Raskar, A. Agrawal, and J. Tumblin, "Coded exposure photography: motion deblurring using fluttered shutter," *ACM Trans. on Graphics* **25**, 795–804 (2006).
- [8] J. Holloway, A. C. Sankaranarayanan, A. Veeraraghavan, and S. Tambe, "Flutter shutter video camera for compressive sensing of videos," in "IEEE Int. Conf. on Computational Photography," (2012), pp. 1–9.
- [9] S. K. Nayar and V. Branzoi, "Adaptive dynamic range imaging: Optical control of pixel exposures over space and time," in "IEEE Int. Conf. on Computer Vision," (2003), pp. 1168–1175.
- [10] X. Lin, J. Suo, G. Wetzstein, Q. Dai, and R. Raskar, "Coded focal stack photography," in "IEEE Int. Conf. on Computational Photography," (2013), pp. 1–9.
- [11] T. Portz, L. Zhang, and H. Jiang, "Random coded sampling for high-speed hdr video," in "IEEE Int. Conf. on Computational Photography," (2015), pp. 1–8.
- [12] G. Bub, M. Tecza, M. Helmes, P. Lee, and P. Kohl, "Temporal pixel multiplexing for simultaneous high-speed, high-resolution imaging," *Nature Methods* **7**, 209–211 (2010).
- [13] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "Flexible voxels for motion-aware videography," in "European Conf. on Computer Vision," (2010), pp. 100–114.
- [14] D. Reddy, A. Veeraraghavan, and R. Chellappa, "P2c2: Programmable pixel compressive camera for high speed imaging," in "IEEE Int. Conf. on Computer Vision and Pattern Recognition," (2011), pp. 329–336.
- [15] Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. K. Nayar, "Video from a single coded exposure photograph using a learned over-complete dictionary," in "IEEE Int. Conf. on Computer Vision," (2011), pp. 287–294.
- [16] D. Liu, J. Gu, Y. Hitomi, M. Gupta, T. Mitsunaga, and S. K. Nayar, "Efficient space-time sampling with pixel-wise coded exposure for high-speed imaging," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **36**, 248–260 (2014).
- [17] J. Gu, Y. Hitomi, T. Mitsunaga, and S. Nayar, "Coded rolling shutter photography: Flexible space-time sampling," in "IEEE Int. Conf. on Computational Photography," (2010), pp. 1–8.
- [18] P. Llull, X. Liao, X. Yuan, J. Yang, D. Kittle, L. Carin, G. Sapiro, and D. J. Brady, "Coded aperture compressive temporal imaging," *Optics Express* **21**, 10526–10545 (2013).
- [19] R. Koller, L. Schmid, N. Matsuda, T. Niederberger, L. Spinoulas, O. Cossairt, G. Schuster, and A. K. Katsaggelos, "High spatio-temporal resolution video with compressed sensing," *Optics Express* **23**, 15992–16007 (2015).

- [20] C. Gao, N. Ahuja, and H. Hua, "Active aperture control and sensor modulation for flexible imaging," in "IEEE Int. Conf. on Computer Vision and Pattern Recognition," (2007), pp. 1–8.
- [21] S. K. Nayar, V. Branzoi, and T. E. Boulton, "Programmable imaging: Towards a flexible camera," *Int. Journal of Computer Vision* **70**, 7–22 (2006).
- [22] H. Mannami, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi, "High dynamic range camera using reflective liquid crystal," in "IEEE Int. Conf. on Computer Vision," (2007), pp. 1–8.
- [23] B. K. Gunturk and M. Feldman, "Frequency division multiplexed imaging," in "IS&T/SPIE Electronic Imaging Conf.," (2013), pp. 86600P–86600P.